	1.0		2.8	2.5
			3.2	2.2
			3.6	
	1.1		4.0	2.0
				1.8
	1.25	1.4		1.6

'OPY RESOLUTION TEST CHART

ARO 19367.18-MA

ON A RANDOM DIFFERENCE EQUATION FOR MATRICES

AND A CHARACTERIZATION OF THE GAMMA DISTRIBUTION

by

Eric S. Tollar

FSU-Statistics Report M676
USARO Technical Report No. D-73

TR-D-73-ARO

October, 1984

The Florida State University
Department of Statistics
Tallahassee, Florida   32306

DTIC
SELECTED
OCT 26 1984

A

Keywords: STOCHASTIC DIFFERENCE EQUATION, RANDOM MATRIX, NORMS SUB-ADDITIVITY, STORAGE LEVEL, GAMMA DISTRIBUTION.

AMS (1980) Subject classifications. Primary 60G99
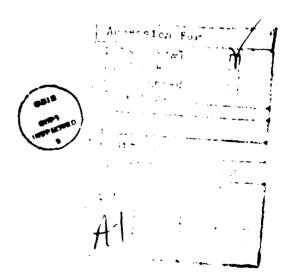
84   10 19 094

On a Random Difference Equation for Matrices

and a Characterization of the Gamma Distribution.

by

Eric S. Tollar

ABSTRACT

The present paper considers the stochastic difference equation $Y_n = M_n Y_{n-1} + Q_n$ where M_n and Q_n are respectively random $d \times d$ matrices and random d-vectors, and obtains some reasonable sufficient conditions on M_n and Q_n under which Y_n converges in distribution. In addition, a particular model is examined when $d = 2$, in which the asymptotic independence of $Y_{1,n}$ and $Y_{2,n}$ results in a characterization of the Gamma distribution.

A-1

1. INTRODUCTION

In this paper we study the limit distribution of the solution Y_n of the difference equation

$$(1.1) \qquad Y_n = M_n Y_{n-1} + Q_n, \quad n \geq 1,$$

where M_n are random $d \times d$ matrices, and Q_n and Y_n are d-vectors (where d could be considered infinite, unless otherwise stated). Throughout we take the sequence of pairs (M_n, Q_n), $n \geq 1$, to be independent and identically distributed. Equation (1.1) first came to our attention in a paper by Bernard, Shenton, and Uppuluri [1], in which it was used as a model for the distribution of radioactive material in the bone structure of humans. Since then, we have seen it arise in a variety of other contexts (see Soloman [9], and Cavalli-Sforza and Feldman [2], for examples).

The asymptotic behavior of (1.1) is examined in [10] by Vervaat in the special case where $d = 1$. A variety of conditions are given for the convergence in distribution of (1.1). In [5], Kesten establishes a reasonably general condition under which (1.1) converges for the cases where $d > 1$. In [5], it is shown that if

$$(1.2) \qquad E(\ln ||M||)^+ < \infty$$

then there is a constant α where

$$(1.3) \qquad \alpha \equiv \lim_{n \to \infty} 1/n \, ||M, \ldots M_n|| \quad \text{a.s.},$$

where for a $d \times d$ matrix M and a d-vector x, we define

$$(1.4) \qquad ||M|| = \max_{|x| = 1} |Mx|, \quad \text{and}$$

$$(1.5) \qquad |x| = \left\{ \sum_{i=1}^{d} x_i^2 \right\}^{\frac{1}{2}}.$$

We then have in [5] that if

$$
(1.6) \qquad \begin{cases} \text{i) } \alpha < 0 \text{ in } (1.3), \text{ and} \\[2mm] \text{ii) there exists a } \beta > 0 \text{ where } E|Q_1|^\beta < \infty, \end{cases}
$$

then (1.1) converges in distribution.

The first objective of this paper is to establish criteria other than (1.6) for the convergence of (1.1). While the conditions established will show that (1.6ii) can be weakened, that is not the aim of this paper. Instead, the criteria attempt to bypass what we perceive to be the major difficulty for applicability of [5], that is, the extreme difficulty in the determination of α in (1.3).

In [7] and [10], the respective authors investigate possible limiting distributions of (1.1) in the case when $d = 1$, and achieve some partial results in classifying the limiting distributions. The second objective of this paper is to examine what criteria on the model (1.1) for $d > 1$ will yield a limit behavior in which the various components of the vector Y_n are asymptotically independent for a special model in which only a one-way reapportioning of material is allowed. We find that in most cases that non-trivial asymptotic independence is in general achieved only if the limit distribution of the components is a gamma distribution. As such, we arrive at a slight generalization of (Lukacs [6]).

2. CONVERGENCE OF THE DIFFERENCE EQUATION.

From (1.1), it is easy to see by iteration that

$$
(2.1) \qquad Y_n = \sum_{i=1}^{n} M_n \cdots M_{i+1} Q_i + M_n \cdots M_1 Y_0,
$$

which, for given Y_0 has the same distribution as

$$
(2.2) \qquad R_n = \sum_{i=1}^{n} M_1 \cdots M_{i-1} Q_i + M_1 \cdots M_n Y_0.
$$

We shall establish that in the case when $M_1 \ldots M_n \to 0$ a.s. exponentially fast, under weak conditions on Y_0 and Q_n, (2.2) will converge almost surely, independent of the distribution of Y_0. Then a variety of conditions which should be reasonably easy to check in particular cases will be given which are sufficient to guarantee that $M_1 \ldots M_n$ converge to zero exponentially fast.

For convenience, for any matrix M and vector Y, we will use the typical notation

(2.3) $\qquad (m_{ij}) \equiv M$, $\qquad (y_i) \equiv Y.$

Also, let us use the following notation, given matrices $M_1, \ldots M_n$, and vector x;

(2.4) $\qquad M^{(n)} = M_1 \ldots M_n,$

(2.5) $\qquad |x|_\infty = \sum_{i=1}^{d} |x_i|,$

(2.6) $\qquad ||M||_\infty = \sup_{|x|_\infty = 1} |Mx|_\infty.$

We now establish the main theorem in this section, which describes sufficient conditions for the almost sure convergence of (2.2).

THEOREM 2.1: If there is a $0 < \lambda < 1$ where either i) $\lambda^{-n} \sup_{1 \le j \le d} (\sum_{i=1}^{d} |m_{i,j}^{(n)}|) \to 0,$

a.s., $E(\ln|Q_1|_\infty)^+ < \infty$, and $|Y_0|_\infty < \infty$, a.s., or ii) $\lambda^{-n} \sup_{1 \le i \le d} (\sum_{j=1}^{d} |m_{i,j}^{(n)}|) \to 0$, a.s.,

$E(\ln(\sup_{1 \le i \le d} |q_{1,i}|)^+) < \infty$, and $\sup_{1 \le i \le d} |y_{0,i}| < \infty$, a.s., then R_n converges almost surely.

PROOF: The proofs that condition i) and condition ii) are sufficient for the almost sure convergence of R_n are similar, and as such, it will only be established for condition i).

We will establish that R_n is a Cauchy sequence with respect to the infinity norm, almost surely. That is, that for all $\epsilon > 0$, there is an N where

(2.7) $\qquad P(\, |R_n - R_m|_\infty < \epsilon, \; \forall m, \, n > N) > 1 - \epsilon.$

We will ignore the term $M_1 \ldots M_n Y_0$, since it is clearly negligible. Then assuming $n > m$,

(2.8)
$$|R_n - R_m|_\infty = \sum_{i=1}^{d} |r_{n,i} - r_{m,i}|$$
$$\leq \sum_{j=m+1}^{n} \sum_{i=1}^{d} \sum_{k=1}^{d} |m_{i,k}^{(j-1)}| |q_{j,k}|$$
$$= \sum_{j=m+1}^{n} \sum_{k=1}^{d} \left(\sum_{i=1}^{d} |m_{i,k}^{(j-1)}| \right) |q_{j,k}|.$$

Let

(2.9)
$$C_N = \{\omega: \sup_{\substack{n \\ n=N \; 1 \leq k \leq d}} \left(\lambda^{-n} \sum_{i=1}^{d} |m_{i,k}^{(n)}| \right) < \epsilon\},$$

we have from (2.8) that

(2.10)
$$|R_n - R_m|_\infty \leq I_{C_m}(\omega) \sum_{j=m+1}^{n} \sum_{k=1}^{d} \epsilon \cdot \lambda^{j-1} |q_{j,k}|$$
$$+ I_{\overline{C}_m}(\omega) \sum_{i=1}^{d} \sum_{k=1}^{d} \left(\sum_{i=1}^{d} |m_{i,k}^{(j-1)} |q_{j,k}| \right),$$

where $I_A(x)$ is the indicator function of a set A.

Thus, from (2.10), we have that

(2.11)
$$P(|R_n - R_m|_\infty > \epsilon, \text{ some } n, m > N)$$
$$\leq P(\overline{C}_N) + P\left(\sum_{j=N+1}^{\infty} \sum_{k=1}^{d} \epsilon \cdot \lambda^{j-1} |q_{j,k}| > \epsilon \right)$$
$$= P(\overline{C}_N) + P\left(\sum_{j=N+1}^{\infty} \epsilon \lambda^{j-1} |Q_j|_\infty > \epsilon \right).$$

By lemma 2.2, to be presented later, we have that $\sum_{i=1}^{\infty} \lambda^{i-1} |Q_i|_{\infty}$ converges almost surely if and only if $E(\ln|Q_1|_{\infty})^{+} < \infty$. From condition i) we have that there exists a λ, $0<\lambda<1$, such that we have that for sufficiently large N, $P(C_n) > 1-\epsilon/2$ for all $n > N$. Thus, it is easy to see from (2.11) that for sufficiently large N,

$$(2.12) \qquad P(|R_n-R_m|_{\infty} > \epsilon, \text{ some } n, m > N) < \epsilon/2 + \epsilon/2 = \epsilon,$$

which establishes the theorem. $\qquad\qquad\qquad\qquad\qquad$ □

In [10], Verwaat establishes the lemma cited above (which is actually more general than the current requirements), which we now state for the sake of completeness.

LEMMA 2.2: For $\{X_n\}$, $\{Y_n\}$, i.i.d. <u>random variables, where</u> $-\infty <E(\ln|X_1|)< 0$, <u>then</u> $\sum_{i=1}^{n} (\prod_{j=1}^{i-1} X_j) Y_i$ <u>converges a.s. if and only if</u> $E(\ln|Y_1|)^{+} < \infty$.

We will now establish several criteria which are sufficient for the exponentially fast convergence of $M_1 M_2 \ldots M_n$. To do this, we will appeal to a general lemma about functions on the matrices.

LEMMA 2.3: <u>If</u> $\{A_n\}$ <u>is a sequence of i.i.d. random matrice, and f is some function such that</u> $f(A_i A_j) \le f(A_i) f(A_j)$ <u>and</u> $E(\ln f(A_1))< 0$, <u>then there is a</u> λ, $0<\lambda<1$ <u>where</u> $P(f(\prod_{i=1}^{n} A_i) > \lambda^n \text{ i.o.}) = 0$.

PROOF: By assumption, we have that

$$(2.13) \qquad \ln f(\prod_{i=1}^{n} A_i) \le \ln(\prod_{i=1}^{n} f(A_i)) = \sum_{i=1}^{n} \ln f(A_i).$$

Thus, if we define $\delta = -E \ln f(A_1)$, we have for any γ where $0<\gamma<\delta$ that

$$(2.14) \qquad P(\sum_{i=1}^{n} \ln f(A_i) > -\gamma n \text{ i.o.}) = 0.$$

As such, from (2.13) we have that

$$(2.15) \qquad P(f(\prod_{i=1}^{n} A_i) > e^{-\gamma n} \text{ i.o.}) \leq P(\sum_{i=1}^{n} \ln f(A_i) > -\gamma n \text{ i.o.}) = 0.$$

We get that for all $\lambda > e^{-\varepsilon}$ the assertion is true. $\qquad \square$

Now we establish the first sufficient criteria for the convergence of R_n which should be reasonably easy to verify in certain cases.

__THEOREM 2.4:__ If $E \ln (||M_1||_\infty) < 0$, $E \ln(|Q_1|_\infty)^+ < \infty$, __and__ $|Y_0|_\infty < \infty$ a.s., __then__ R_n converges almost surely.

__PROOF:__ By condition i) of Theorem 2.1, we need only establish that
$\lambda^{-n} \sup_{1 \leq j \leq d} \left(\sum_{i=1}^{d} |m_{i,j}^{(n)}| \right) \to 0$, a.s. This follows quickly from Lemma 2.2 by noting
that for all x where $|Bx|_\infty \neq 0$,

$$(2.16) \qquad |ABx|_\infty = \frac{|ABx|_\infty}{|Bx|_\infty} |Bx|_\infty \leq ||A||_\infty |Bx|_\infty,$$

where the inequality follows from the definition of theorem (2.6). Thus we have that

$||AB||_\infty \leq ||A||_\infty ||B||_\infty$. As such, from the assumptions and lemma 2.2 we have that

$$(2.17) \qquad P(||\prod_{i=1}^{n} M_i||_\infty > \lambda^n \text{ i.o.}) = 0.$$

This yields that for $\lambda < \lambda_1 < 1$, we have that

$$(2.18) \qquad \lambda_1^{-n} ||\prod_{i=1}^{n} M_i||_\infty \to 0, \text{ a.s. }.$$

As can be easily established from (2.6), we have that

$$(2.19) \qquad ||\prod_{i=1}^{n} M_i||_\infty = \sup_{1 \leq j \leq d} (\sum_{i=1}^{d} |m_{i,j}^{(n)}|),$$

and the theorem is completed. $\qquad \square$

While the condition $E \ln||M_1||_\infty < 0$ is not as formidable a condition to check as at first it appears, since formula (2.19) gives some hope of verifying $E \ln||M_1||_\infty < 0$, nonetheless it is perhaps too difficult for certain models. The following theorem gives a simpler condition to check.

Using the matrix notation $E|M| = (E|m_{i,j}|)$, we have the following theorem.

THEOREM 2.5: If $d < \infty, ||E|M_1| \; ||_\infty < 1$, $E \ln(|Q_1|_\infty)^+ < \infty$, and $|Y_0|_\infty < \infty$, a.s., then R_n converges almost surely.

PROOF: Once again, by condition i) of theorem 2.1, we need only establish that for $0 < \lambda < 1$, $\lambda^{-n} \sup_{1 \leq j \leq d} (\sum_{i=1}^{d} |m_{i,j}^{(n)}|) \to 0$, a.s. Since $||E|M^{(1)}|||_\infty < 1$, we can choose a λ and λ_1 where $||E|M^{(1)}| \; ||_\infty < \lambda_1 < \lambda < 1$. From (2.16), we have that

$$(2.20) \qquad || \prod_{k=1}^{n} E|M_k| \; ||_\infty \leq \prod_{k=1}^{n} ||E|M_k| \; ||_\infty < \lambda_1^{n}.$$

Also, we have from the independence of the matrices

$$(2.21) \qquad E|m_{i,j}^{(2)}| = E|\sum_{k=1}^{d} m_{1,i,k} \; m_{2,k,j} | \leq \sum_{k=1}^{d} E|m_{1,i,k}|E|m_{2,k,j}|,$$

which along with (2.19) and (2.20) yields

$$(2.22) \qquad ||E|M_1 \ldots M_n| \; ||_\infty \leq || \prod_{k=1}^{n} E|M_k| \; ||_\infty < \lambda_1^{n}.$$

Also, we have for $\lambda > \lambda_1$ that

$$(2.23) \qquad P(\cup_{n=m+1}^{\infty} \lambda^{-n} \sup_{1 \leq j \leq d} (\sum_{i=1}^{d} |m_{i,j}^{(n)}|) \leq \sum_{n=m+1}^{\infty} P(\lambda^{-n} \sup_{1 \leq k \leq d} (\sum_{i=1}^{d} |m_{i,j}^{(n)}|) > \epsilon).$$

$$\leq \sum_{n=m+1}^{\infty} \lambda^{-n} \epsilon^{-1} E \sup_{1 \leq j \leq d} (\sum_{i=1}^{d} |m_{i,j}^{(n)}|).$$

Since

$$(2.24) \quad E \sup_{1 \leq j \leq d} \left(\sum_{i=1}^{d} |m_{i,j}^{(n)}| \right) \leq \sum_{j=1}^{d} \left(E \sum_{i=1}^{d} |m_{i,j}^{(n)}| \right) \leq d \sup_{1 \leq j \leq d} \left(\sum_{i=1}^{d} E|m_{i,j}^{(n)}| \right)$$

$$= d \, \| E|M_1 \ldots M_n| \, \|_\infty,$$

we get from (2.23) that

$$(2.25) \quad P \left(\bigcup_{n=m+1}^{\infty} \lambda^{-n} \sup_{1 \leq j \leq d} \left(\sum_{i=1}^{d} |m_{i,j}^{(n)}| \right) > \epsilon \right) \leq \sum_{n=m+1}^{\infty} \lambda^{-1} \epsilon^{-1} d \| E|M_1 \ldots M_n| \|_\infty$$

$$\leq \sum_{n=m+1}^{\infty} d \, \epsilon^{-1} \left(\frac{\lambda_1}{\lambda} \right)^n = d\epsilon^{-1} \left(\frac{\lambda_1}{\lambda} \right)^{m+1} \left(1 - \frac{\lambda_1}{\lambda} \right)^{-1}$$

where the second inequality follows from (2.22).

As such, for m sufficiently large, we have

$$(2.26) \quad P \left(\bigcup_{n=m+1}^{\infty} \lambda^{-n} \sup_{1 \leq j \leq d} \left(\sum_{i=1}^{d} |m_{i,j}^{(n)}| \right) > \epsilon \right) < \epsilon,$$

which completes the proof. $\square$

We now introduce another functional on matrices which is very similar to the spectral radius, first introduced by Dobrushin [3] in the case of stochastic matrices. For any $d \times d$ matrix $P = (p_{ij})$, let us define an auxiliary matrix $\hat{P} = (\hat{p}_{ij})$ by

$$(2.27) \quad \hat{p}_{ij} = \begin{cases} 1 & \text{if } i=1, \, j=1 \\ 0 & \text{if } i > 1, \, j=1 \\ 1 - \sum_{i=1}^{d} p_{ij} & \text{if } i=1, \, j > 1 \\ p_{i-1 \, j-1} & \text{if } i > 1, \, j > 1 \end{cases}$$

Then we define our functional $\delta(P)$ by

$$(2.28) \quad \delta(P) = \sup_{1 \leq i,k \leq d+1} \left(\sum_{j=1}^{d+1} \left[\hat{p}_{ji} - \hat{p}_{jk} \right]^* \right).$$

LEMMA 2.6: For any $d \times d$ matrices P, Q, where

$$\sup_{1 \leq j \leq d} \left(\sum_{i=1}^{d} |p_{ij}| \right) < \infty, \quad \sup_{1 \leq j \leq d} \left(\sum_{i=1}^{d} |q_{ij}| \right) < \infty, \quad \delta(PQ) < \delta(P) \, \delta(Q).$$

PROOF: The functional δ is very similar to that of Dobrushin [3]. An examination of the proof that for stochastic matrices P', Q', we have $\delta(P'Q') \leq \delta(P')\delta(Q')$, (see, Isaacson and Madsen [4]), yields that the crucial properties are

(2.29)

$$\text{i)} \quad \sum_{j=1}^{d+1} \hat{p}_{ji} = 1, \quad \text{for all } i$$

$$\text{ii)} \quad \sum_{\ell \in E} \sum_{j=1}^{d+1} \hat{p}_{\ell j} (\hat{q}_{ji} - \hat{q}_{jk}) = \sum_{j=1}^{d+1} \sum_{\ell \in E} \hat{p}_{\ell j}(\hat{q}_{ji} - \hat{q}_{jk}) \text{ all } i, k.$$

Since both are satisfied by the definition of $\hat{P}$, $\hat{Q}$, the proof is complete. □

It should be noted that $\delta(\cdot)$ and $||\cdot||_\infty$ are very similar. In fact, if P is a non-negative matrix, then it is easy to show that $\delta(P) = ||P||_\infty$. However, if P is allowed to have negative values, it is possible that $\delta(P) < ||P||_\infty$ or $||P||_\infty < \delta(P)$. Because of this similarity in behavior, it is easy to see that a theorem similar to theorem 2.4 can be established for $\delta(M_1)$ by methods identical to those used in theorem 2.4, (we need only notice that from (2.28) we have $\delta(P) \geq \sup_{1 \leq i \leq d} (\sum_{j=1}^{d} p_{ji} +)$, $\delta(P) \geq \sup_{1 \leq i \leq d} (\sum_{j=1}^{d} p_{ji}^-)$, which yields that $2\delta(P) \geq \sup_{1 \leq i \leq d} (\sum_{j=1}^{d} |p_{ij}|))$. As such, we state without proof the following theorem.

Theorem 2.7. If $E \ln\delta(M_1) < 0$, $E \ln(|Q_1|_\infty)^+ < \infty$, and $|Y_0|_\infty < \infty$ a.s., then R_n converges almost surely.

It should be noted that theorems 2.4, 2.5, and 2.7 were applications of condition i) of theorem 2.1. It can be easily seen that by considering M_1', the transpose of M_1, that theorems similar to 2.4, 2.5 and 2.7 can be established using condition ii) and virtually identical proofs to the theorems already established instead. However, both the statements and proofs of these theorems will be omitted in this paper.

3. LIMITING INDEPENDENCE FOR A PARTICULAR RANDOM DIFFERENCE MODEL.

In [7] and [10], the respective authors investigate possible limiting distributions of (1.1) for the case of d=1, and achieve some partial results in classifying possible limiting distributions. For d=2, we will examine for a special model the conditions under which the two components are asymptotically independent. The model to be considered is the difference equation

$$(3.1) \quad \begin{pmatrix} Y_{1,n} \\ Y_{2,n} \end{pmatrix} = \begin{pmatrix} V_n & 0 \\ 1-V_n & W_n \end{pmatrix} \begin{pmatrix} Y_{1,n-1} \\ Y_{2,n-1} \end{pmatrix} + \begin{pmatrix} Q_n \\ 0 \end{pmatrix},$$

where $\{V_n, W_n, Q_n\}$ is an i.i.d. sequence, and V_n, W_n, Q_n are independent of each other for all n. This represents a one-way flow storage model, in which at step n new material is added to component one via Q_n, material is transferred from component one to component two via $1-V_n$, and material is lost from the system from component two via $1-W_n$.

For the model given in (3.1), if we let $Y_n \overset{d}{\to} Y$, and let

$$(3.2) \quad \begin{cases} \phi(s,t) = E(e^{isY_1 + itY_2}) \\ \psi(s) = Ee^{isQ_1}, \end{cases}$$

then it is easy to verify from (3.1) that

$$(3.3) \quad \phi(s,t) = \psi(s)E\phi(sV_1 + t(1-V_1), tW_1).$$

Since Y_1 and Y_2 are independent if and only if $\phi(s,t) = \phi(s,0) \phi(0,t)$, we have from (3.3) that

$$(3.4) \quad \phi(s,t) = \psi(s) E\phi(sV_1 + t(1-V_1),0) E\phi(0,tW_1).$$

Also from (3.3) it can also be shown that

(3.5) $\phi(s,t) = \psi(s) \ E\phi(sV_1,0) \ E\phi(t(1-V_1),tW_1)$

$= \psi(s) \ E\phi(sV_1,0) \ E\phi(t(1-V_1),0) \ E\phi(0,tW_1).$

If we let

(3.6) $A = \{s: \ \psi(s) = 0 \text{ or } E\phi(0,sW_1) = 0\},$

and if A^C is dense, then by equating (3.4) and (3.5) one can see that for all $(s,t) \in A^C \times A^C$ that

(3.7) $E\phi(sV_1 + t(1-V_1),0) = E\phi(sV_1,0) \ E\phi(t(1-V_1),0).$

Then, by continuity of characteristic functions, we get that (3.7) holds for all $(s,t) \in R \times R$. Thus, under conditions sufficient to guarantee that A^C is dense, we get from (3.7) that Y_1, Y_2 are independent if and only if for V_1 independent of Y_1, V_1Y_1 and $(1-V_1)Y_1$ are independent.

In the following, we will say

(3.8) $X \sim \Gamma(\lambda,\beta)$ if $P(X \leq x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \int_0^x \Gamma(\beta)^{-1}\lambda(\lambda y)^{\beta-1}e^{-\lambda y}dy & \text{if } x > 0, \end{cases}$

and

(3.9) $X \sim B(\alpha,\beta)$ if $P(X \leq x) = \begin{cases} 0 & \text{if } x \leq 0 \text{ or } x \geq 1 \\ \int_0^x B(\alpha,\beta)^{-1} y^{\alpha-1}(1-y)^{\beta-1}dy & \text{if } 0 < x < 1. \end{cases}$

The main result of the section is the following theorem.

THEOREM 3.1:   For model (3.1), if V_1 and Y_1 are independent then V_1Y_1 and $(1-V_1)Y_1$ are independent if and only if one of the following six conditions are true:

1) $Y_1 \equiv 0$

2) $Y_1 \equiv c, \ V_1 \equiv d$

3) $V_1 \equiv 0$

4) $V_1 \equiv 1$

5) $Y_1 \sim \Gamma(\lambda, \alpha+\beta), \ V_1 \sim B(\alpha, \beta)$

6) $-Y_1 \sim \Gamma(\lambda, \alpha+\beta), \ V_1 \sim B(\alpha, \beta)$

To establish theorem 3.1, we will first establish the following lemma and intermediate theorem.

LEMMA 3.2: Let U, W be independent random variables, where $U > W > 0$. Then $U(U-W)^{-1}$ and $U - W$ are independent if and only if $U \equiv c$, $W \equiv d$, $c > d > 0$.

PROOF: It is clear there must be a constant e where

$$(3.10) \quad P(U > e) = P(W \le e) = 1$$

If not, then we could find a b where $P(U \le b) > 0$, $P(W > b) > 0$, but $P(U \le b, W > b) = 0$, a contradiction.

Let

$$(3.11) \quad \begin{cases} b_2 = \inf \{b: \ P(W \le b) = 1\}, \\ b_3 = \sup \{b: \ P(U \ge b) = 1\}. \end{cases}$$

Then $0 < W \le b_2 \le b_3 \le U$.

Since $U - W \ge b_3 - b_2$, $W \le b_2$, we have that

$$(3.12) \quad U(U - W)^{-1} \le \frac{b_3}{b_3 - b_2}.$$

Also, it is clear $P(U - W < b_3 - b_2 + \epsilon) > 0$, for all $\epsilon > 0$.

Since $U(U - W)^{-1}$, $U - W$ are independent, we have that

$$(3.13) \quad P(U(U - W)^{-1} > b_3(b_3-b_2+\epsilon)^{-1}) = P(U(U - W)^{-1} > b_3(b_3-b_2+\epsilon)^{-1} | U - W < b_3-b_2+\epsilon).$$

And $U - W < b_3-b_2+\epsilon$ implies

$$(3.14) \quad U < b_3+\epsilon \text{ and } W > b_2-\epsilon.$$

As such, we have that $U - W < b_3-b_2+\epsilon$ implies

$$(3.15) \quad U(U - W)^{-1} > 1 + \frac{b_2-\epsilon}{b_3-b_2+\epsilon} = \frac{b_3}{b_3-b_2+\epsilon} .$$

Thus $P(U(U - W)^{-1} > \frac{b_3}{b_3-b_2+\epsilon} | U - W < b_3-b_2+\epsilon) = 1$ for all $\epsilon > 0$.

As such, from (3.13), we get

$$(3.16) \quad P(U(U - W)^{-1} \geq \frac{b_3}{b_3-b_2}) = 1.$$

Combining (3.12) and (3.16), we get $P(U(U - W)^{-1} = \frac{b_3}{b_3-b_2}) = 1$

which in turn implies $U = b_3$, $V = b_2$. $\quad\square$

Using this lemma, we can establish the next intermediate theorem.

__THEOREM 3.3:__ If $Y_1 \geq 0$, and if V_1, Y_1 are independent, then V_1Y_1 and $(1-V_1)Y_1$ are independent if and only if one of the five conditions below are true:

1) $Y_1 \equiv 0$,

2) $V_1 \equiv 0$,

3) $V_1 \equiv 1$,

4) $Y_1 \equiv c$, $V_1 \equiv d$,

5) $Y_1 \sim \Gamma(\lambda,\alpha+\beta)$, $V \sim B(\alpha,\beta)$.

PROOF: The sufficiency is obvious, so we need only establish the necessity of the conditions. For convenience, let

(3.15) $X = V_1 Y_1$, $Z = (1-V_1)Y_1$.

Further, let

(3.16)
$$
\begin{cases}
A_1 = \{Y_1 = 0\} = \{X = 0, \ Z = 0\} \\
A_2 = \{Y_1 > 0, \ V_1 < 0\} = \{X < 0, \ Z > 0\} \\
A_3 = \{Y_1 > 0, \ V_1 = 0\} = \{X = 0, \ Z > 0\} \\
A_4 = \{Y_1 > 0, \ 0 < V_1 < 1\} = \{X > 0, \ Z > 0\} \\
A_5 = \{Y_1 > 0, \ V_1 = 1\} = \{X > 0, \ Z = 0\} \\
A_6 = \{Y_1 > 0, \ V_1 > 1\} = \{X > 0, \ Z < 0\}.
\end{cases}
$$

Also, for the random variables V_1, Y_1, X, Z, we define, respectively, the variable V_A, Y_A, X_A, Z_A for any arbitrary set A as the restriction of the variable to the set A. That is, the distribution of variable V_A is given by

(3.17) $P(V_A \leq x) = [P(V \epsilon A)]^{-1} P(V \leq x, \ V \epsilon A)$

with the other variables defined identically.

As can be easily established, if V_1 and Y_1 are independent, than V_{A_i} and Y_{A_i} are also independent for all i. Similarly, if X and Z are independent then so are X_{A_i} and Z_{A_i} for all i.

To prove the theorem, we examine three cases,

1) $P(Y_1=0) = 1$ 2) $P(Y_1=0) < 1$, $P(0<V_1<1) > 0$ and 3) $P(Y_1=0) < 1$, $P(0<V_1<1) = 0$

CASE I: $P(Y_1=0) = 1$.

In this case, we have nothing to prove, this being condition 1) of the theorem.

CASE II: $P(Y_1=0) < 1$, $P(0<V_1<1) > 0$.

In this case, $P(A_4) = P(Y_1>0) \cdot P(0<V_1<1) > 0$.

Clearly $X_{A_4} > 0$, $Z_{A_4} > 0$, and X_{A_4} and Z_{A_4} are independent. Also V_{A_4}, Y_{A_4} are independent, and as can be seen from (3.15), we have

$$X_{A_4} + Z_{A_4} = Y_{A_4}, \ X_{A_4}(X_{A_4} + Y_{A_4})^{-1} = V_{A_4} .$$

Thus from Lukacs characterization of the gamma distribution (see [6]), we have that

(3.18) $\qquad X_{A_4} \sim \Gamma(\lambda,\alpha), \ Z_{A_4} \sim \Gamma(\lambda,\beta)$

Also, we have that

(3.19) $\qquad Y_{A_4} = X_{A_4} + Z_{A_4} \sim \Gamma(\lambda,\alpha+\beta).$

Since $Y_{\{Y>0\}} \sim Y_{A_i}$ for $i \neq 1$ (from the independence of Y_1 and V_1), we have for any $i \neq 1$ where $P(A_i) > 0$, that

(3.20) $\qquad Y_{A_i} \sim Y_{A_4}$ for $i \neq 1$.

By similar arguments, we can also establish that for any i where $P(A_i) > 0$, we must have
$$X_{A_i} \sim X_{A_4} \text{ for } i > 4,$$

(3.21)

$$Z_{A_i} \sim Z_{A_4} \text{ for } 1<i<4.$$

Let us assume $P(A_2) > 0$. Then by letting

(3.22) $\qquad U = Z_{A_2}, \ W = -X_{A_2},$

we have that $U > W > 0$. As can be easily established, U, W are independent, and $W(U-W)^{-1}$, $U-W$ are independent. Also, since $W(U-W)^{-1} = U(U-W)^{-1} -1$, we can apply Lemma 3.2, yielding $Z_{A_2} \equiv c$, $X_{A_2} \equiv -d$. But since $Z_{A_4} \sim \Gamma(\lambda,\beta)$, and from (3.21) we have $Z_{A_2} \sim Z_{A_4}$, we have a contradiction which yields that $P(A_2) = 0$.

If $P(A_6) > 0$, we can generate by similar methods a contradiction for the distribution of X_{A_6}, so $P(A_6) = 0$. If $P(A_3) > 0$, we have $X_{A_3} = 0$, so $Y_{A_3} = Z_{A_3}$.

But from (3.20) and (3.21), we have $Y_{A_3} \sim \Gamma(\lambda,\beta)$ and $Z_{A_3} \sim \Gamma(\lambda,\alpha+\beta)$ yielding by contradiction that $P(A_3) = 0$. Similarly, we can show that $P(A_5) = 0$, yielding that $P(0<V_1<1) = 1$. Thus we have that, letting $p_1 = P(Y_1 = 0)$, that

(3.23) $P(X \leq x, Z \leq z)$

$$= p_1 \; I(X=0, Z=0) + (1-p_1) \; \Gamma(\lambda,\alpha)(x) \; \Gamma(\lambda,\beta)(z)$$

Since $P(X=0) = P(Z=0) = P(X=0, Z=0) = p_1$, from independence of X and Z, we have that $p_1 = p_1^2$ which yields that $p_1 = 0$ since $p_1 < 1$. This yields condition 5) of the theorem.

CASE III: $P(Y_1=0) < 1$, $P(0<V_1<1) = 0$.

Since $P(0<V_1<1) = 0$, we have $P(X>0, Z>0) = 0$. Thus, either $P(X>0) = 0$, or $P(Z>0) = 0$. Assume $P(Z>0) = 0$. Then $P(V_1 \leq 0) = 1$. In addition, assume that $P(V_1<0) > 0$, so we have $P(A_2) > 0$, and by applying Lemma 3.2 to X_{A_2}, Z_{A_2}, we get $Z_{A_2} = c$, $X_{A_2} = -d$, $c>d>0$. Thus $Y_{A_2} = Z_{A_2} + X_{A_2} = c-d>0$. If $P(V_1=0) > 0$ also, we have $P(A_3) > 0$. By (3.20) and (3.21), we get that $Y_{A_3} \sim Y_{A_2} = c - d$, and $Z_{A_3} \sim Z_{A_2} = c$. But we must have $Z_{A_3} = Y_{A_3}$, which gives a contradiction. Thus, if $P(V_1 < 0) > 0$, then $P(V_1 = 0) = 0$ which yields that $P(V_1 < 0) = 1$.

Thus, we have for $P(Y_1 = 0) = p_1$

(3.24) $\begin{cases} P(X=0, Z=0) = p_1 \\[2mm] P(X=-d, Z=-c) = (1-p_1) \; , \end{cases}$

which clearly contradicts the independence of X and Z unless $p_1 = 0$ or $p_1 = 1$. Thus we have X and Z are degenerate, which yields condition 4).

If we assume instead that $P(V_1=0) > 0$ in addition to $P(Z>0) = 0$, a contradiction similar to the one above will show that $P(V_1=0) = 1$, which is condition 2).

If we assume that $P(X>0) = 0$, arguments identical to those above will show that $P(V_1=1) = 1$ or $P(V_1>1) = 1$, which yield, respectively, condition 4) or condition 3). □

We are now ready to establish the main theorem of this section.

THEOREM 3.4. If V_1, Y_1 are not both degenerate, then V_1, Y_1 independent and V_1Y_1 and $(1-V_1)Y_1$ are independent if and only if one of the five conditions below are true:

1) $Y_1 \equiv 0$,

2) $V_1 \equiv 0$,

3) $V_1 \equiv 1$,

4) $Y_1 \sim \Gamma(\lambda,\alpha+\beta)$, $V_1 \sim B(\alpha,\beta)$

5) $-Y_1 \sim \Gamma(\lambda,\alpha+\beta)$, $V_1 \sim B(\alpha,\beta)$.

PROOF: Again the sufficiency is obvious, so we need only establish the necessity of the conditions. If we have either $P(Y_1 \geq 0) = 1$, or $P(Y_1 \leq 0) = 1$, then by theorem 3.3 we have that one of the 5 conditions must hold (where the new condition, condition 5) follow when $P(Y_1 \leq 0) = 1$). Thus from now on we will assume $P(Y_1<0) > 0$, $P(Y_1>0) > 0$.

In addition to the sets A_1 to A_6 defined in theorem 3.3, we define the sets

$$(3.25) \quad \begin{cases} A_7 = \{Y_1<0, V_1<0\} = \{X>0, Z<0, X+Z<0\} \\ A_8 = \{Y_1<0, V_1=0\} = \{X=0, Z<0\} \\ A_9 = \{Y_1<0, 0<V_1<1\} = \{X<0, Z<0\} \\ A_{10} = \{Y_1<0, V_1=1\} = \{X<0, Z=0\} \\ A_{11} = \{Y_1<0, V_1>1\} = \{X<0, Z>0, X+Z<0\}, \end{cases}$$

and we redefine

$$(3.26) \quad \begin{cases} A_2 = \{X<0, \ Z>0, \ X+Z>0\} \\ A_6 \quad \{X>0, \ Z<0, \ X+Z>0\} \end{cases}$$

Note that while for all i, A_i can be expressed as $A_i = \{Y_1 \epsilon A, \ V_1 \epsilon B\}$ (for some borel sets A, B) the sets A_2, A_6, A_7, A_{11}, cannot be expressed as sets of the form $\{X \epsilon A, \ Z \epsilon B\}$.

Let us assume in addition to $P(Y_1<0) > 0$, $P(Y_1>0) > 0$, that $P(0<V_1<1) > 0$ (this assumption will be shown to yield a contradiction later).

Thus we have $P(A_4) > 0$, $P(A_9) > 0$. By appealing to Lukacs characterization as in Theorem 3.3, we get

$$(3.27) \quad \begin{cases} X_{A_4} \sim \Gamma(\lambda_1,\alpha_1), \ Z_{A_4} \sim \Gamma(\lambda_1,\beta_1), \ Y_{A_4} \sim \Gamma(\lambda_1,\alpha_1+\beta_1) \\ -X_{A_9} \sim \Gamma(\lambda_2,\alpha_2), \ -Z_{A_9} \sim \Gamma(\lambda_2,\beta_2), \ -Y_{A_9} \sim \Gamma(\lambda_1,\alpha_1+\beta). \end{cases}$$

By arguments similar to those used for (3.20) and (3.21), we have that if $P(A_6 \cup A_7) > 0$, $P(A_2 \cup A_{11}) > 0$, then

$$(3.28) \quad \begin{cases} X_{A_6 \cup A_7}, \ -Z_{A_6 \cup A_7}, \ \sim \Gamma(\lambda_1,\alpha_1) \times \Gamma(\lambda_2,\beta_2) \\ -X_{A_2 \cup A_{11}} \quad Z_{A_2 \cup A_{11}} \sim \Gamma(\lambda_2,\alpha_2) \times \Gamma(\lambda_1,\beta_1). \end{cases}$$

Since $P(A_6 \cup A_7) > 0$, assume that $P(A_6) > 0$. Thus we get for $x_0 > 0$, $y_0 < 0$ that

$$(3.29) \quad P(X_{A_6} \le x_0, \ Z_{A_6} \le z_0) = \iint_A \frac{\lambda_1^{\alpha_1} x^{\alpha_1-1} e^{-\lambda_1 x} \lambda_2^{\beta_2}(-Z)^{\beta_2-1} e^{\lambda_2 Z}}{\Gamma(\alpha_1)\Gamma(\beta_2)P(A_6)} dx dz$$

where $A = \{(x,z) : x<x_0, \ z \le z_0, \ x+z>0\}$.

Also, by arguments similar to those used for (3.20) and (3.21), we have that

$$(3.29) \quad Y_{A_6} = X_{A_6} + Z_{A_6} \sim \Gamma(\lambda_1,\alpha_1+\beta_1).$$

Thus, by transformation of variables in (3.28), we get that

$$(3.30) \quad \lambda_1^{\beta_1} y^{\alpha_1 + \beta_1 - 1} = \frac{\Gamma(\alpha_1 + \beta_1) \lambda_2^{\beta_2}}{\Gamma(\alpha_1)\Gamma(\beta_1)P(A_6)} \int_0^\infty (y+w)^{\alpha_1 - 1} e^{-(\lambda_1 + \lambda_2)w} w^{\beta_2 - 1} dw.$$

By examining the behavior of the right hand side of (3.30) and the left hand side of (3.30) (particularly as y approaches 0), we get that equality is impossible for $\alpha_1 > 0$, $\beta_1 > 0$. Thus, we have that $P(A_6) = 0$. Similarly, we can establish that $P(A_7) = 0$. Thus $P(X > 0, Z < 0) = P(A_6 \cup A_7) = 0$. However, since we have assumed that $P(Y_1 > 0) > 0$, $P(Y_1 < 0) > 0$, and $P(0 < V_1 < 1) > 0$, we get from the definition of X and Z (see (3.15)), that $P(X > 0) > 0$, $P(Z < 0) > 0$. This yields a contradiction in the assumption $P(0 < V_1 < 1) > 0$. Thus, we conclude that $P(0 < V_1 < 1) = 0$, and we have that $P(X > 0, Z > 0) = P(X < 0, Z < 0) = 0$.

From this, we can quickly deduce the rest of the criteria. Since $P(Y_1 > 0) > 0$, and $P(Y_1 < 0) > 0$, then the assumption that $P(V_1 < 0)$ yields from (3.15) that $P(X < 0) > 0$, $P(Z > 0) > 0$. Independence of X and Z yields that $P(X < 0, Z < 0) > 0$, which is a contradiction. Thus, $P(V_1 < 0) = 0$. In a similar manner, $P(V_1 > 1) = 0$. Thus $P(V_1 = 0) + P(V_1 = 1) = 1$. If $P(V_1 = 0) > 0$, $P(V_1 = 1) > 0$, we again can show from (3.15) that $P(X > 0) > 0$, $P(Z > 0) > 0$, which again contradicts $P(X > 0, Z > 0) = 0$. Thus we have either $P(V_1 = 0) = 0$, $P(V_1 = 1) = 1$ (condition 3) or $P(V_1 = 0) = 1$, $P(V_1 = 1) = 0$ (condition 2). $\square$

To relate Theorem 3.4 to the asymptotic independence of $Y_{1,n}$ and $Y_{2,n}$, we only need to observe as remarked after (3.7) that for any condition sufficient to guarantee the density of A^c (see (3.6)), Y_1 and Y_2 are independent if and only if for V_1, Y_1 independent, $V_1 Y_1$ and $(1-V_1)Y_1$ are also independent. As such, for A^c dense, Theorem 3.4 gives the necessary and sufficient condition for asymptotic independence. The most reasonable and realistic condition on $\{Q_n, V_n, W_n\}$ for Theorem 3.4 to be applicable are given in the following corollary.

COROLLARY 3.5   If $Q_n \geq 0$, $0 \leq V_n \leq 1$, $W_n \geq 0$ then $Y_{1,n}$ and $Y_{2,n}$ are asymptotically independent if and only if Q_n has a Gamma distribution and V_n has a Beta distribution, or one of the four trivial conditions of Theorem 3.3 are met by Q_n and V_n.

The proof follows immediately from Theorem 3.4 and the fact that the characteristic functions of non-negative random variables have dense support (see Smith [8]).

Clearly, other restrictions to Q_n, V_n, and W_n will yield that the only case of asymptotic independence of $Y_{1,n}$ and $Y_{2,n}$ are when Q_n has a Gamma distribution and V_n has a Beta distribution. However, the more interesting question of a characterization of the distributions of Q_n, V_n and W_n that result in $Y_{1,n}$ and $Y_{2,n}$ being asymptotically independent appears to be an open question.

REFERENCES

[1] Bernard, S.R., Shenton, L.R., and Uppuluri, V.R.R. (1967). Stochastic models for the distribution of radioactive material in a connected system of compartments. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.*, 481-510.

[2] Cavalli-Sforza, L. and Feldman, M.W. (1973). Models for cultural inheritance I. Group mean and within group variation. *Theoret. Popn. Biol.*, 4, 42-55.

[3] Dobrushin, R.L, (1956). The central limit theorem for nonhomogeneous Markov chains. I-II. *Teor. Veroyatnost, i. Primenen.*, 1, 72-88, 365-425.

[4] Isaacson, D.L., and Madsen, R.W. (1976). *Markov Chains, Theory and Applications*, J. Wiley and Sons, New York.

[5] Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices, *Acta. Math.* 131, 207-248.

[6] Lukacs, E. (1955). A characterization of the gamma distribution. *Ann. Math. Statist.*, 26, 319-324

[7] Paulson, A.S. and Uppuluri, V.R.R. (1972). Limit laws on a sequence determined by a random difference equation governing a one-compartment system. *Math. Biosci.*, 13, 325-333.

[8] Smith, W.L. (1962). A note on characteristic functions which vanish identically in an interval, *Proc. Camb. Phil. Soc. Math. Phys. Sci.*, 58, 430-432.

[9] Solomon, F. (1975). Random walks in a random environment. *Ann. Prob.*, 3, 1-31.

[10] Verwaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Prob.*, 11, 750-783.

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER FSU 676 USARO D-73	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and subtitle) ON A RANDOM DIFFERENCE EQUATION FOR MATRICES AND A CHARACTERIZATION OF THE GAMMA DISTRIBUTION		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Eric S. Tollar		8. CONTRACT OR GRANT NUMBER(s) USARO DAAG 29-82-K-0168
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL 32306		10. PROGRAM ELEMENT, PROJECT, TASK AREA, & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office-Durham P.O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE October, 1984
		13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this report)

Distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report.

18. SUPPLEMENTARY NOTES

19. KEY WORDS

STOCHASTIC DIFFERENCE EQUATION; RANDOM MATRIX; NORMS SUB-ADDITIVITY; STORAGE LEVEL; GAMMA DISTRIBUTION.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The present paper considers the stochastic difference equation $Y_n = M_n Y_{n-1} + Q_n$ where M_n and Q_n are respectively random $d \times d$ matrices and random d-vectors, and obtains some reasonable sufficient conditions on M_n and Q_n under which Y_n converges in distribution. In addition, a particular model is examined when $d = 2$, in which the asymptotic independence of $Y_{1,n}$ and $Y_{2,n}$ results in a characterization of the Gamma distribution.

END

FILMED

DTIC